

PAGE: 1

**RAW SEQUENCE LISTING
PATENT APPLICATION US/08/864,955**DATE: 03/16/98
TIME: 17:05:50**INPUT SET: S24135.raw**

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING**1 (1) General Information:****5 (i) APPLICANT: Cold Spring Harbor Laboratory****7 (ii) TITLE OF INVENTION: Novel Human cdc25 Genes, Encoded
8 Products and Uses Thereof****10 (iii) NUMBER OF SEQUENCES: 31****12 (iv) CORRESPONDENCE ADDRESS:****13 (A) ADDRESSEE: FOLEY, HOAG & ELIOT LLP
14 (B) STREET: ONE POST OFFICE SQUARE
15 (C) CITY: Boston
16 (D) STATE: Massachusetts
17 (E) COUNTRY: U.S.A.
18 (F) ZIP: 02109****20 (v) COMPUTER READABLE FORM:****21 (A) MEDIUM TYPE: Floppy disk
22 (B) COMPUTER: IBM PC compatible
23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
24 (D) SOFTWARE: ASCII(text)****26 (vi) CURRENT APPLICATION DATA:****27 (A) APPLICATION NUMBER: 08/864,955
28 (B) FILING DATE: 29 MAY 1997
29 (C) CLASSIFICATION:****31 (viii) ATTORNEY/AGENT INFORMATION:****32 (A) NAME: Vincent, Matthew P.
33 (B) REGISTRATION NUMBER: 36,709
34 (C) REFERENCE/DOCKET NUMBER: MIV-019.05****36 (ix) TELECOMMUNICATION INFORMATION:****37 (A) TELEPHONE: 617-832-1242
38 (B) TELEFAX: 617-832-7000****41 (2) INFORMATION FOR SEQ ID NO:1:****43 (i) SEQUENCE CHARACTERISTICS:****44 (A) LENGTH: 2419 base pairs
45 (B) TYPE: nucleic acid
46 (C) STRANDEDNESS: double****ENTERED**

**RAW SEQUENCE LISTING
PATENT APPLICATION US/08/864,955**

DATE: 03/16/98
TIME: 17:05:54

INPUT SET: S24135.raw

```

47          (D) TOPOLOGY: linear
48
49          (ii) MOLECULE TYPE: DNA (genomic)
50
51          (ix) FEATURE:
52              (A) NAME/KEY: CDS
53              (B) LOCATION: 460..2031
54
55          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
56
57 CGAAAGGCCG GCCTTGGCTG CGACAGCCTG GGTAAGAGGT GTAGGTCGGC TTGGTTTTCT      60
58
59 GCTACCCCGA GCTGGGCAAG CGGGTTGGGA GAACAGCGAA GACAGCGTGA GCCTGGCCG      120
60
61 TTGCCTCGAG GCTCTCGCCC GGCTTCTCTT GCCGACCCGC CACGTTGTT TGGATTTAAT      180
62
63 CTTACAGCTG GTTGCCGGCG CCCGCCGCC CGCTGGCCTC GCGGTGTGAG AGGGAAGCAC      240
64
65 CCGTGCCTGT GGCTGGTGGC TGGCGCCTGG AGGGTCCGCA CACCCGCCG GCCGCCGCC      300
66
67 TTTGCCCGCG GCAGCCGCGT CCCTGAACCG CGGAGTCGTG TTTGTGTTG ACCCGCGGGC      360
68
69 GCCGGTGGCG CGCGGCCGAG GCCGGTGTGCG GCGGGGGCGGG GCGGTGCGGG CGGAGGCAGA      420
70
71 GGAAGAGGGA GCGGGAGCTC TGCGAGGCCG GCGGCCGCC ATG GAA CTG GGC CCG      474
72                               Met Glu Leu Gly Pro
73                               1           5
74
75 AGC CCC GCA CCG CGC CGC CTG CTC TTC GCC TGC AGC CCC CCT CCC GCG      522
76 Ser Pro Ala Pro Arg Arg Leu Leu Phe Ala Cys Ser Pro Pro Pro Ala
77                               10          15          20
78
79 TCG CAG CCC GTC GTG AAG GCG CTA TTT GGC GCT TCA GCC GCC GGG GGA      570
80 Ser Gln Pro Val Val Lys Ala Leu Phe Gly Ala Ser Ala Ala Gly Gly
81                               25          30          35
82
83 CTG TCG CCT GTC ACC AAC CTG ACC GTC ACT ATG GAC CAG CTG CAG GGT      618
84 Leu Ser Pro Val Thr Asn Leu Thr Val Thr Met Asp Gln Leu Gln Gly
85                               40          45          50
86
87 CTG GGC AGT GAT TAT GAG CAA CCA CTG GAG GTG AAG AAC AAC AGT AAT      666
88 Leu Gly Ser Asp Tyr Glu Gln Pro Leu Glu Val Lys Asn Asn Ser Asn
89                               55          60          65
90
91 CTG CAG AGA ATG GGC TCC TCC GAG TCA ACA GAT TCA GGT TTC TGT CTA      714
92 Leu Gln Arg Met Gly Ser Ser Glu Ser Thr Asp Ser Gly Phe Cys Leu
93                               70          75          80          85
94
95 GAT TCT CCT GGG CCA TTG GAC AGT AAA GAA AAC CTT GAA AAT CCT ATG      762
96 Asp Ser Pro Gly Pro Leu Asp Ser Lys Glu Asn Leu Glu Asn Pro Met
97                               90          95          100
98
99 AGA AGA ATA CAT TCC CTA CCT CAA AAG CTG TTG GGA TGT AGT CCA GCT      810

```

**RAW SEQUENCE LISTING
PATENT APPLICATION US/08/864,955**

 DATE: 03/16/98
 TIME: 17:05:57
INPUT SET: S24135.raw

100	Arg Arg Ile His Ser Leu Pro Gln Lys Leu Leu Gly Cys Ser Pro Ala			
101	105	110	115	
102				
103	CTG AAG AGG AGC CAT TCT GAT TCT CTT GAC CAT GAC ATC TTT CAG CTC	858		
104	Leu Lys Arg Ser His Ser Asp Ser Leu Asp His Asp Ile Phe Gln Leu			
105	120	125	130	
106				
107	ATC GAC CCA GAT GAG AAC AAG GAA AAT GAA GCC TTT GAG TTT AAG AAG	906		
108	Ile Asp Pro Asp Glu Asn Lys Glu Asn Glu Ala Phe Glu Phe Lys Lys			
109	135	140	145	
110				
111	CCA GTA AGA CCT GTA TCT CGT GGC TGC CTG CAC TCT CAT GGA CTC CAG	954		
112	Pro Val Arg Pro Val Ser Arg Gly Cys Leu His Ser His Gly Leu Gln			
113	150	155	160	165
114				
115	GAG GGT AAA GAT CTC TTC ACA CAG AGG CAG AAC TCT GCC CAG CTC GGA	1002		
116	Glu Gly Lys Asp Leu Phe Thr Gln Arg Gln Asn Ser Ala Gln Leu Gly			
117	170	175	180	
118				
119	ATG CTT TCC TCA AAT GAA AGA GAT AGC AGT GAA CCA GGG AAT TTC ATT	1050		
120	Met Leu Ser Ser Asn Glu Arg Asp Ser Ser Glu Pro Gly Asn Phe Ile			
121	185	190	195	
122				
123	CCT CTT TTT ACA CCC CAG TCA CCT GTG ACA GCC ACT TTG TCT GAT GAG	1098		
124	Pro Leu Phe Thr Pro Gln Ser Pro Val Thr Ala Thr Leu Ser Asp Glu			
125	200	205	210	
126				
127	GAT GAT GGC TTC GTG GAC CTT CTC GAT GGA GAG AAT CTG AAG AAT GAG	1146		
128	Asp Asp Gly Phe Val Asp Leu Leu Asp Gly Glu Asn Leu Lys Asn Glu			
129	215	220	225	
130				
131	GAG GAG ACC CCC TCG TGC ATG GCA AGC CTC TGG ACA GCT CCT CTC GTC	1194		
132	Glu Glu Thr Pro Ser Cys Met Ala Ser Leu Trp Thr Ala Pro Leu Val			
133	230	235	240	245
134				
135	ATG AGA ACT ACA AAC CTT GAC AAC CGA TGC AAG CTG TTT GAC TCC CCT	1242		
136	Met Arg Thr Thr Asn Leu Asp Asn Arg Cys Lys Leu Phe Asp Ser Pro			
137	250	255	260	
138				
139	TCC CTG TGT AGC TCC AGC ACT CGG TCA GTG TTG AAG AGA CCA GAA CGT	1290		
140	Ser Leu Cys Ser Ser Thr Arg Ser Val Leu Lys Arg Pro Glu Arg			
141	265	270	275	
142				
143	TCT CAA GAG GAG TCT CCA CCT GGA AGT ACA AAG AGG AGG AAG AGC ATG	1338		
144	Ser Gln Glu Ser Pro Pro Gly Ser Thr Lys Arg Arg Lys Ser Met			
145	280	285	290	
146				
147	TCT GGG GCC AGC CCC AAA GAG TCA ACT AAT CCA GAG AAG GCC CAT GAG	1386		
148	Ser Gly Ala Ser Pro Lys Glu Ser Thr Asn Pro Glu Lys Ala His Glu			
149	295	300	305	
150				
151	ACT CTT CAT CAG TCT TTA TCC CTG GCA TCT TCC CCC AAA GGA ACC ATT	1434		
152	Thr Leu His Gln Ser Leu Ser Ala Ser Ser Pro Lys Gly Thr Ile			

**RAW SEQUENCE LISTING
PATENT APPLICATION US/08/864,955**

 DATE: 03/16/98
 TIME: 17:06:01
INPUT SET: S24135.raw

153	310	315	320	325	
154					
155	GAG AAC ATT TTG GAC AAT GAC CCA AGG GAC CTT ATA GGA GAC TTC TCC				1482
156	Glu Asn Ile Leu Asp Asn Asp Pro Arg Asp Leu Ile Gly Asp Phe Ser				
157	330	335	340		
158					
159	AAG GGT TAT CTC TTT CAT ACA GTT GCT GGG AAA CAT CAG GAT TTA AAA				1530
160	Lys Gly Tyr Leu Phe His Thr Val Ala Gly Lys His Gln Asp Leu Lys				
161	345	350	355		
162					
163	TAC ATC TCT CCA GAA ATT ATG GCA TCT GTT TTG AAT GGC AAG TTT GCC				1578
164	Tyr Ile Ser Pro Glu Ile Met Ala Ser Val Leu Asn Gly Lys Phe Ala				
165	360	365	370		
166					
167	AAC CTC ATT AAA GAG TTT GTT ATC ATC GAC TGT CGA TAC CCA TAT GAA				1626
168	Asn Leu Ile Lys Glu Phe Val Ile Ile Asp Cys Arg Tyr Pro Tyr Glu				
169	375	380	385		
170					
171	TAC GAG GGA GGC CAC ATC AAG GGT GCA GTG AAC TTG CAC ATG GAA GAA				1674
172	Tyr Glu Gly Gly His Ile Lys Glu Ala Val Asn Leu His Met Glu Glu				
173	390	395	400	405	
174					
175	GAG GTT GAA GAC TTC TTA TTG AAG AAG CCC ATT GTA CCT ACT GAT GGC				1722
176	Glu Val Glu Asp Phe Leu Leu Lys Lys Pro Ile Val Pro Thr Asp Gly				
177	410	415	420		
178					
179	AAG CGT GTC ATT GTT GTG TTT CAC TGC GAG TTT TCT TCT GAG AGA GGT				1770
180	Lys Arg Val Ile Val Val Phe His Cys Glu Phe Ser Ser Glu Arg Gly				
181	425	430	435		
182					
183	CCC CGC ATG TGC CGG TAT GTG AGA GAG AGA GAT CGC CTG GGT AAT GAA				1818
184	Pro Arg Met Cys Arg Tyr Val Arg Glu Arg Asp Arg Leu Gly Asn Glu				
185	440	445	450		
186					
187	TAC CCC AAA CTC CAC TAC CCT GAG CTG TAT GTC CTG AAG GGG GGA TAC				1866
188	Tyr Pro Lys Leu His Tyr Pro Glu Leu Tyr Val Leu Lys Gly Gly Tyr				
189	455	460	465		
190					
191	AAG GAG TTC TTT ATG AAA TGC CAG TCT TAC TGT GAG CCC CCT AGC TAC				1914
192	Lys Glu Phe Phe Met Lys Cys Gln Ser Tyr Cys Glu Pro Pro Ser Tyr				
193	470	475	480	485	
194					
195	CGG CCC ATG CAC CAC GAG GAC TTT AAA GAA GAC CTG AAG AAG TTC CGC				1962
196	Arg Pro Met His His Glu Asp Phe Lys Glu Asp Leu Lys Lys Phe Arg				
197	490	495	500		
198					
199	ACC AAG AGC CGG ACC TGG GCA GGG GAG AAG AGC AAG AGG GAG ATG TAC				2010
200	Thr Lys Ser Arg Thr Trp Ala Gly Glu Lys Ser Lys Arg Glu Met Tyr				
201	505	510	515		
202					
203	AGT CGT CTG AAG AAG CTC TGAGGGCGGC AGGACCAGCC AGCAGCAGCC				2058
204	Ser Arg Leu Lys Lys Leu				
205	520				

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/864,955DATE: 03/16/98
TIME: 17:06:04

INPUT SET: S24135.raw

206 CAAGCTTCCC TCCATCCCC TTTACCTCT TTCCTGCAGA GAAACTTAAG CAAAGGGAC 2118
207
208 AGCTGTGTGA CATTGGAGA GGGGCCCTGG GACTTCATG CCTTAAACCT ACCTCCCACA 2178
209
210 CTCCAAGGT TGGAGCCAG GGCATCTTGC TGGCTACGCC TCTTCTGTCC CTGTTAGACG 2238
211
212 TCCTCCGTCC ATATCAGAAC TGTGCCACAA TGCAGTTCTG AGCACCGTGT CAAGCTGCTC 2298
213
214 TGAGCCACAG TGGGATGAAC CAGCCGGGGC CTTATCGGGC TCCAGCATCT CATGAGGGGA 2358
215
216 GAGGAGACGG AGGGGAGTAG AGAAGTTAC ACAGAAATGC TGCTGGCCAA ATAGCAAAGA 2418
217
218
219 G 2419
220
221
222 (2) INFORMATION FOR SEQ ID NO:2:
223
224 (i) SEQUENCE CHARACTERISTICS:
225 (A) LENGTH: 523 amino acids
226 (B) TYPE: amino acid
227 (D) TOPOLOGY: linear
228
229 (ii) MOLECULE TYPE: protein
230
231 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
232
233 Met Glu Leu Gly Pro Ser Pro Ala Pro Arg Arg Leu Leu Phe Ala Cys
234 1 5 10 15
235
236 Ser Pro Pro Pro Ala Ser Gln Pro Val Val Lys Ala Leu Phe Gly Ala
237 20 25 30
238
239 Ser Ala Ala Gly Gly Leu Ser Pro Val Thr Asn Leu Thr Val Thr Met
240 35 40 45
241
242 Asp Gln Leu Gln Gly Leu Gly Ser Asp Tyr Glu Gln Pro Leu Glu Val
243 50 55 60
244
245 Lys Asn Asn Ser Asn Leu Gln Arg Met Gly Ser Ser Glu Ser Thr Asp
246 65 70 75 80
247
248 Ser Gly Phe Cys Leu Asp Ser Pro Gly Pro Leu Asp Ser Lys Glu Asn
249 85 90 95
250
251 Leu Glu Asn Pro Met Arg Arg Ile His Ser Leu Pro Gln Lys Leu Leu
252 100 105 110
253
254 Gly Cys Ser Pro Ala Leu Lys Arg Ser His Ser Asp Ser Leu Asp His
255 115 120 125
256
257 Asp Ile Phe Gln Leu Ile Asp Pro Asp Glu Asn Lys Glu Asn Glu Ala
258 130 135 140

INPUT SET: S24135.raw******* PREVIOUSLY ERRORED SEQUENCES - EDITED *******

1403 (2) INFORMATION FOR SEQ ID NO:31:
1404
1405 (i) SEQUENCE CHARACTERISTICS:
1406 (A) LENGTH: 11 amino acids
1407 (B) TYPE: amino acid
1408 (D) TOPOLOGY: linear
1409
1410 (ii) MOLECULE TYPE: peptide
1411
1412 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:
1413
1414 Leu Val Phe His Cys Glu Xaa Xaa Xaa Xaa Arg
1415 1 5 10
1416
1417
1418
1419
1420

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/864,955

DATE: 03/16/98
TIME: 17:06:09

INPUT SET: S24135.raw

Line

Error

Original Text